

TEACHER'S GUIDE

DNA *to Darwin Case Study*

The origins and evolution of HIV

Version 1.3

Anne Fischer

Formerly of the Max Planck
Institute for Evolutionary
Anthropology, Leipzig

Dean Madden (Ed.)

NCBE, University of Reading



Charles Darwin

The origins and evolution of HIV

In this Case Study the origin of HIV-1 is investigated by comparing genetic sequences from Human Immunodeficiency Virus (HIV) with Simian Immunodeficiency Virus (SIV). DNA sequences derived from the *pol* gene of viral RNA sequences are used for this purpose.

When did HIV evolve?

When the first cases of AIDS were reported in the early 1980s, the cause of the disease was a mystery. The discovery of HIV in 1983* led to considerable speculation about the origin of the virus. Many ideas were suggested, some of them fanciful, including a curse from Tutenkhamen's tomb and the deliberate creation of the virus by genetic modification.

Recent genetic research shows that the truth is less dramatic. Most estimates suggest that HIV evolved from SIV in the 1930s, although a 2008 study placed the origin of HIV between 1884 and 1924 (Worobey, et al., 2008). The earliest known case of AIDS occurred in 1959 in the Belgian Congo (today, the 'Democratic Republic of Congo'). Several cases have been confirmed from the 1960s and '70s. By the 1980s, international travel coupled with other factors (such as the increased use of blood transfusion, intravenous drug abuse and sexual promiscuity) had allowed the virus to spread from Africa to the USA, Europe and rest of the world.

* although the virus was discovered in 1983, the link to AIDS was made in 1984.

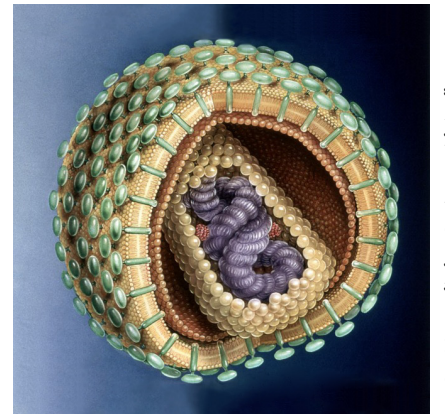
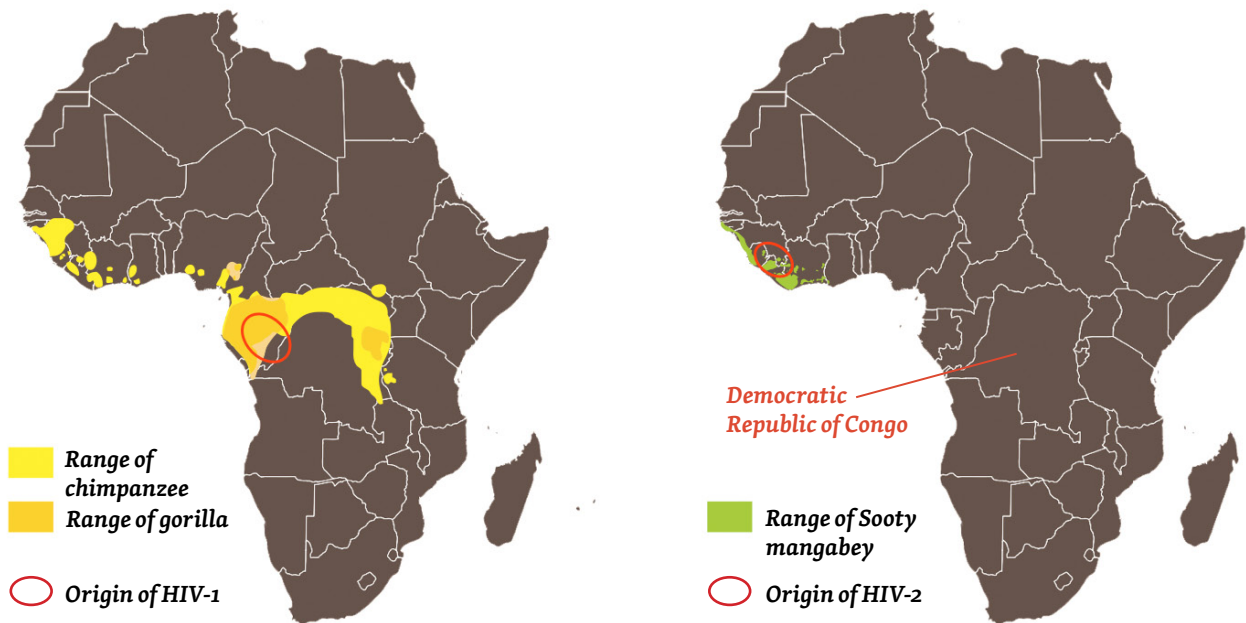


IMAGE FROM: Medical Art Service, Munich / Wellcome Images

HIV, the virus that causes AIDS, was identified in 1983. It has evolved from similar viruses (SIVs) that affect monkeys and apes.



The red rings on the maps above show the approximate areas in which the two forms of HIV are thought to have originated. These areas coincide with the distribution of primates carrying SIV, which is genetically similar to HIV.

Two types of HIV

The virus that is the primary cause of AIDS worldwide is now called HIV-1. In 1986, a second type of HIV, called HIV-2, was isolated from AIDS patients in West Africa. HIV-1 and HIV-2 have the same modes of transmission and are associated with similar opportunistic infections when a patient's immune system is compromised. HIV-2 is initially less virulent and is confined largely to countries in West Africa e.g., Senegal, Liberia, Ghana, and the Ivory Coast.

The geographical range of HIV-2 coincides with the historical habitat of the Sooty mangabey (*Cercocebus atys atys*) (Santiago *et al.*, 2005). These monkeys were often kept as pets or hunted for food, so the SIV could have entered the human population through cuts and scratches from animals.

Origin of HIV-1

The origin of HIV-1 has proved more elusive. In 1999, an international group of researchers announced that they had found a SIVcpz that was almost identical to HIV-1 (Gao *et al.*, 1999). The SIVcpz came from a frozen sample taken from a captive chimpanzee (subspecies *Pan troglodytes troglodytes*). Later, it became possible to extract SIVcpz from faecal samples collected from wild chimpanzees, and this confirmed that they were indeed a reservoir of SIVcpz (Keele *et al.*, 2006). Later that year, members of the same team suggested that one form of HIV-1 (Group O) might be derived from gorilla SIV (SIVgor) (Van Heuverswyn *et al.*, 2006).

The student activity focuses on HIV-1 and the three strains of the virus that have been defined: M, N and O. More than 90% of HIV infections worldwide are of the HIV-1 M group (M stands for 'main'); Group O is restricted to west-central Africa and Group N is very rare and is found only in Cameroon. Each of these groups probably represents a separate transfer of SIV from apes into the human population. (Similarly, there are several groups of HIV-2 which are thought to derive from separate 'jumps' of SIVs into humans.)

A new subgroup of HIV-1, P, has recently been proposed (Plantier *et al.*, 2009), and it is thought that this strain was derived from a gorilla SIV.

Origin of SIVcpz

A few years after they had proposed that SIVcpz was the source of HIV-1, some of the researchers mentioned above suggested that SIVcpz was a hybrid of two monkey SIVs (Bailes *et al.*, 2003). Their suggestion was based on an analysis of the sequences of the viruses. They proposed that a chimp had become infected simultaneously with two monkey SIVs which had recombined within the host to form a new virus. Since chimpanzees are known to hunt and kill monkeys for food, transmission of the viruses to chimps could have occurred in a similar manner to that in which SIVs are thought to have transferred to humans.

HIV denialism

The belief that HIV is not the cause of AIDS is has proved remarkably resilient, even in the face of overwhelming scientific and medical evidence to the contrary. HIV denialism started in the USA, but its impact has been greatest in South Africa. For years, officials in South Africa denied the link between HIV and AIDS. This has contributed that country's poor response to its AIDS epidemic. It is estimated that between 2000 and 2005 the denialist policy of Thabo Mbeki's government led to the early deaths of more than 330,000 South Africans (Chigwedere *et al.*, 2008). The current South African government has pledged to adopt more effective policies to stem the spread of AIDS and to treat the disease. HIV denialism is mentioned in the Students' materials as it helps in part to explain the high incidence of AIDS in South Africa.

Other theories

Several theories have been proposed over the last 25 years to explain the origin and spread of HIV. These include the idea that it was spread by a trial of oral polio vaccine in the 1950s, or that it was deliberately created in the laboratory. Although many of these ideas are fanciful, several have been suggested by scientists. Students may encounter these ideas, particularly while searching the Web or reading older literature. The AVERT website has details of the main ideas and the evidence to support or refute them: www.avert.org

General reading

The making of the fittest. DNA and the ultimate forensic record of evolution by Sean B. Carroll (2009) Quercus Books (Paperback) ISBN: 978 1847247247.
A popular lay account of some of the molecular evidence for evolution.

Reading the story in DNA: A beginner's guide to molecular evolution by Lindell Bromham (2008) Oxford University Press (Paperback) ISBN: 978 0199290918.

An engaging textbook on molecular evolution, which assumes no specialist mathematical knowledge and takes the reader from first principles. Although it's aimed at undergraduates, this superb book contains sufficient detail for PhD students, yet parts will appeal equally to 16–19 year-olds. One of the case studies in the book is the (now discredited) Oral Polio Vaccine (CHAT) theory of the spread of HIV-1 (See: Worobey *et al.*, 2004).

References

- Gao, F. *et al.* (1999) Origin of HIV-1 in the chimpanzee *Pan troglodytes troglodytes*. *Nature* **397**: 436–441.
- Bailes, E. *et al.* (2003) Hybrid origin of SIV in chimpanzees. *Science*, **300**: 1713.

- Worobey, M. *et al.* (2004) Origin of AIDS: Contaminated polio vaccine theory refuted. *Nature* **428**: 820.
- Santiago, M.L. *et al.* (2005) Simian immunodeficiency virus infection in free-ranging Sooty Mangabeys (*Cercocebus atys atys*) from the Taï Forest, Côte d'Ivoire: Implications for the origin of epidemic human immunodeficiency virus Type 2. *Journal of Virology* **79**: 12515–12527.
- Keele, B. F. *et al.* (2006) Chimpanzee reservoirs of pandemic and nonpandemic HIV-1. *Science* **313**: 523–526.
- Van Heuverswyn, F. *et al.* (2006) Human immunodeficiency viruses: SIV infection in wild gorillas. *Nature* **444**: 164.
- Worobey, M. *et al.* (2008) Direct evidence of extensive diversity of HIV-1 in Kinshasa by 1960. *Nature* **455**: 661–664.
- Chigwedere P. *et al.* (2008) Estimating the lost benefits of antiretroviral drug use in South Africa. *Journal of Acquired Immune Deficiency Syndrome*. **49**: 410–5.
- Plantier, J-C. *et al.* (2009) A new human immunodeficiency virus derived from gorillas. *Nature Medicine* **15**: 871–872.

Requirements

Software

The software required, *Geneious*, can be downloaded free-of-charge from: www.geneious.com. The software is available for Windows, Macintosh and Linux operating systems. Only the free, 'basic' version of the software is required for this activity.

DNA sequence data

Students will need the *Geneious* document containing 16 aligned DNA sequences derived from HIV and SIV: **DNA-HIV1andSIV.geneious**

Students' worksheets

Students will require copies of worksheets. Pages 1–7 are an introduction to the exercise and can be set as preparatory work; pages 8–14 give step-by-step instructions for carrying out the computer analysis.

Presentations

Teachers may find the *QuickTime* animation and *PowerPoint* or *Keynote* presentations helpful for introducing this exercise. *QuickTime* may be downloaded free-of-charge from the Apple web site: www.apple.com/quicktime.

Educational aims

The activity introduces HIV and AIDS. The aim of the analysis is to use nucleotide sequence data from the *pol* gene of HIV-1 and SIV sequenced in chimpanzees and gorillas and see how closely related these sequences are. The activity reinforces students' understanding of DNA structure and function including the genetic code and protein synthesis. It also introduces the principle of sequence alignment.

There are three parts to the exercise:

- translation of the nucleotide sequences into protein sequences;
- alignment of the sequences, which means that the amino acids will be aligned according to degree of similarity;
- building a phylogenetic tree, which shows the degree of relatedness between HIV and SIV sequences. The tree is a graphical representation of the evolutionary relationships between the virus sequences.

Possible extension activities include downloading and comparing the genetic sequences of Sooty mangabey SIV (SIVmm) and HIV-2.

Sequences for these additional species may be downloaded from the *GenBank* database and aligned by the students, or they can be provided to the students ready-aligned.

Prerequisite knowledge

Students will need to understand the structure and function of DNA, including the base-pairing mechanism and how DNA encodes amino acids. They will need to be taught the principle of sequence alignment.

The introductory information on HIV and AIDS could be provided as preparation before the lesson in which the computer analysis is done.

Other useful resources

The Genius of Charles Darwin (2008) Channel 4

Region 2, PAL: one DVD. 4DVD. ASIN: B001B8NPL6

Three documentaries presented by Richard Dawkins. 3 x 48 minutes.

Individual programmes from this series are available for purchase from iTunes in the UK: www.apple.com/uk/itunes/

Part of Episode 1, 'Viruses', features HIV/AIDS in Africa, how the virus is still evolving and resistance to HIV in the human population. A clip from the programme can be used either to introduce this activity or to conclude it.

Why creationism is wrong and evolution is right

Lecture given by Steve Jones at The Royal Society, 11 April 2006

Available at: <http://royalsociety.org/Event.aspx?id=1666>

The title of Steve Jones's talk is misleading, as he barely mentions creationism. He starts by using the evolution of languages as an analogy for biological evolution, then discusses the evolution of HIV and resistance to the virus.

Arkive

Photographs and video clips of the animal species mentioned in this case study may be viewed at: www.arkive.org

Avert

Avert is a UK-based AIDS charity: www.avert.org

Answers to the questions on the worksheets

Page 2

- a. HIV/AIDS is commonest in Southern, East and West-Central Africa.

Page 5

- b. One explanation for the distribution of HIV/AIDS might be that the virus originated there. Others include population, healthcare and cultural differences, leading to greater spread of the virus, or the possibility of a genetic susceptibility to the virus amongst indigenous populations (note the high incidence in Guyana and the Caribbean, which have populations of African descent).

Page 9

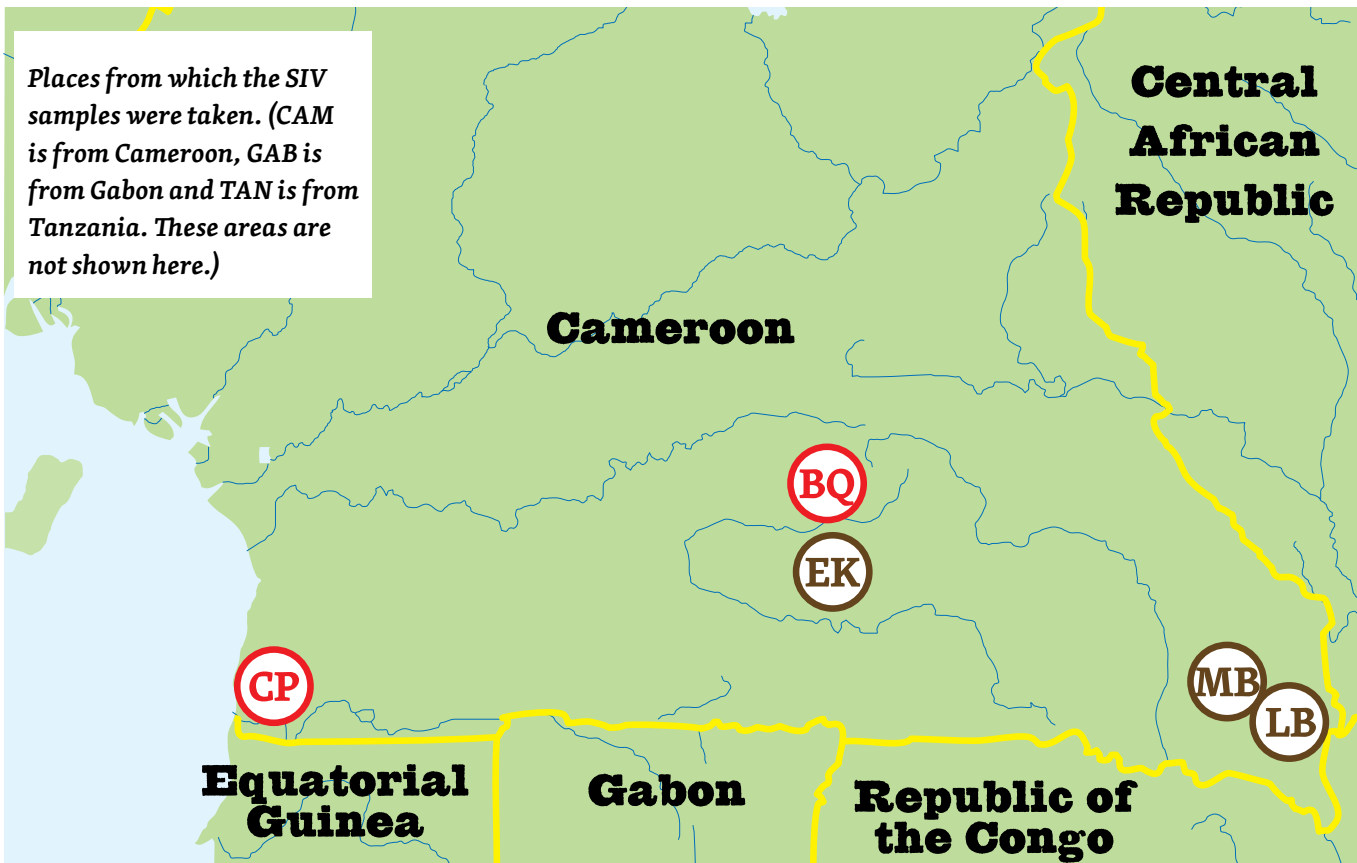
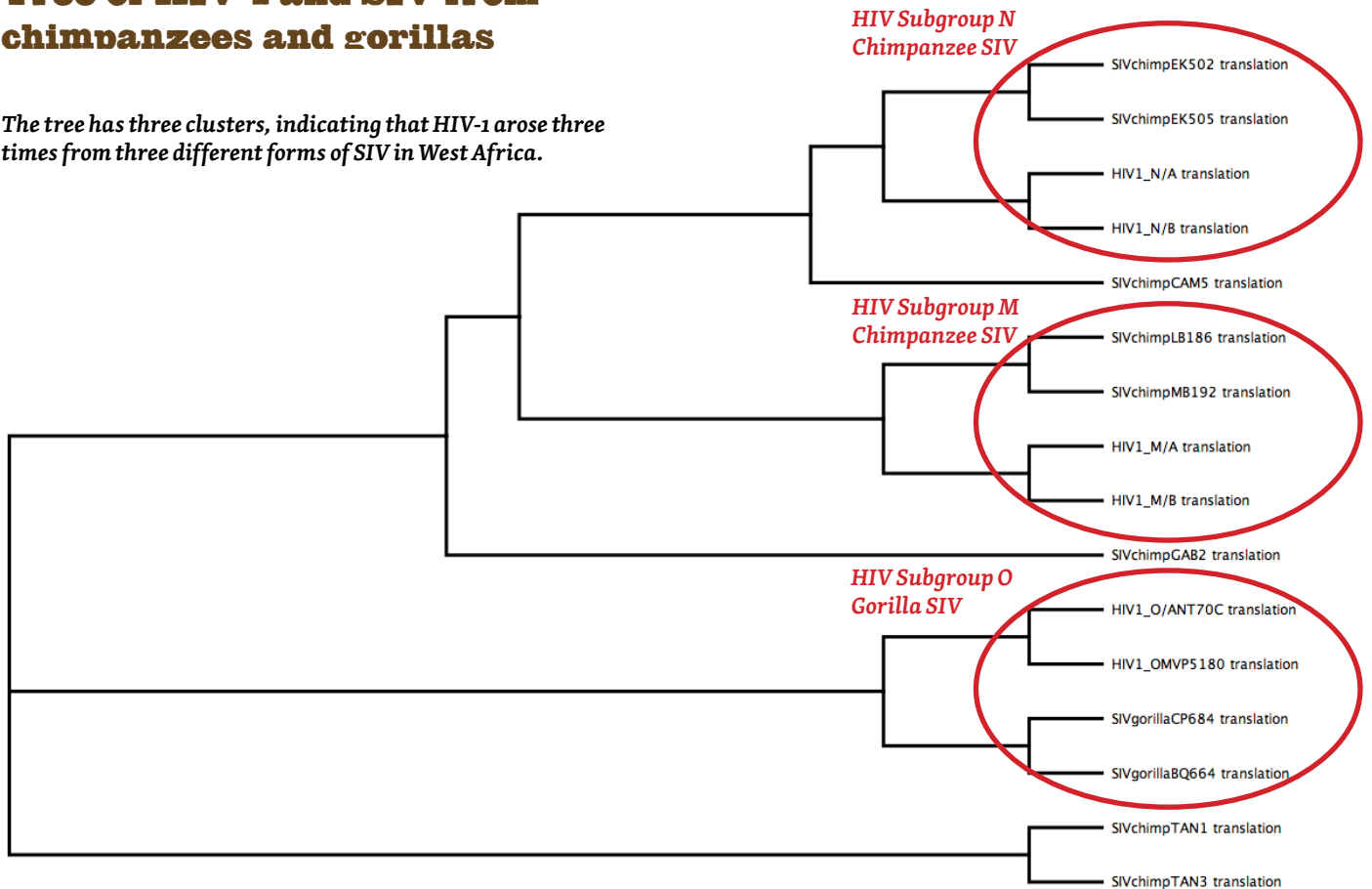
- c. The base thymine (T) is present in DNA and not in RNA, where it is replaced by uracil (U).
- d. Retroviruses have RNA as their genetic material.
- e. Students might reasonably say that the RNA sequence data has been transcribed to DNA. In fact, the sequences are RNA in which the uracil (U) has been replaced by thymine (T). Hence they look like DNA, but they are RNA, which is stored in *GenBank* in this way.

Page 14

- f. All the HIV sequences are obviously from humans; the SIV chimp and gorilla sequences are also clearly named. The codes merely reflect different sampling locations.
- g. There are three main clusters of HIV and SIV sequences. There are six human sequences: two from subgroup M; two from N and two from O. These pairs of sequences are distant from one another, due to mutations occurring in the virus. SIVTAN₁ and SIVTAN₃ are the most distant sequences from all the others. These sequences were found in chimpanzees from eastern Africa, whereas all the other chimpanzee SIV sequences are from central Africa.
- h. Each subgroup of HIV-1 is more closely related to a different primate virus and not to another HIV-1 subgroup.
- i. The fact that each subgroup of HIV-1 is more closely related to different primate virus and not to another HIV-1 subgroup implies that at least three independent transmissions must have occurred. Two occurred from chimpanzee to human and one from gorilla to human.
- j. There is a geographic coincidence between the natural habitat of chimpanzees and gorillas and the epicentres of the HIV epidemics, although the greatest incidence of HIV in Southern Africa, probably because the high population density there permits its rapid spread. Central Africa, in the region encompassing Gabon, Cameroon, Equatorial Guinea and the Republic of Congo is the only place where all three subgroups of HIV-1 are found. The most closely-related sequences of chimpanzee SIV and gorilla SIV have been found in chimpanzees and gorillas inhabiting the same areas.

Tree of HIV-1 and SIV from chimpanzees and gorillas

The tree has three clusters, indicating that HIV-1 arose three times from three different forms of SIV in West Africa.



AFTER: Van Heuverswyn, F. et al. (2006)